Yale University
EliScholar – A Digital Platform for Scholarly Publishing at Yale

Cowles Foundation Discussion Papers Cowles Foundation

3-1-1966

Simulation of Socio-Economic Systems. Part I: General Considerations

Martin Shubik

Follow this and additional works at: https://elischolar.library.yale.edu/cowles-discussion-paper-series

Part of the Economics Commons

Recommended Citation
Shubik, Martin, "Simulation of Socio-Economic Systems. Part I: General Considerations" (1966). *Cowles Foundation Discussion Papers*. 433.
https://elischolar.library.yale.edu/cowles-discussion-paper-series/433

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 203

Note: Cowles Foundation Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. Requests for single copies of a Paper will be filled by the Cowles Foundation within the limits of the supply. References in publications to Discussion Papers (other than mere acknowledgment by a writer that he has access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

SIMULATION OF SOCIO-ECONOMIC SYSTEMS

PART I: GENERAL CONSIDERATIONS

Martin Shubik

March 1, 1966

SIMULATION OF SOCIO-ECONOMIC SYSTEMS*

PART I: GENERAL CONSIDERATIONS

Martin Shubik

---

*Research undertaken by the Cowles Commission for Research in Economics under Contract Nonr-3055(00) with the Office of Naval Research.

1. Introduction

In this paper a general discussion is presented of some of the basic considerations which must be given prior to applying the techniques of simulation to problems of economic development.

1.1. Mathematical Models and Simulations

Two types of mathematical models are familiar to most economists interested in growth and international trade. They are sequential models with relations formulated in terms of difference or differential equations; and models involving the simultaneous evaluation of several variables given a set of relationships between the variables.

A simple example of the first type might consist of three or four highly aggregate accounting and behavioral equations in which, as in the example below the whole economy is represented by two or three sectors and each sector is represented by a single behavioral relation meant to typify all actions of that sector.

$$Y_t = G_t + I_t + C_t \qquad \text{(Accounting relation)}$$

$$C_t = \alpha Y_{t-1} \qquad \text{(Behavior relation)}$$

$$I_t = \beta (C_t - C_{t-1}) \qquad \text{(Behavior relation)}$$

$$G_t = k \qquad \left(\begin{array}{c}\text{Behavior relation} \\ \text{-- Policy Control --}\end{array}\right)$$

where

Y_t = Gross National Product in time t

G_t = Government spending in time t

C_t = Consumption in time t

I_t = Private investment in time t

Given that care has been taken to specify the ranges of definition of these relationships, in a case as simple as the one given here we could solve the system mathematically with little difficulty and (if we believed that it were a reasonably good representation of reality) we could directly examine the effect of changes in the parameters.

In this model a limited amount of exploration of government policy can be made by observing the effect of different size k's. Suppose however we wished to consider, for example, a counter-cyclical policy related say to the level of unemployment and the growth rate of GNP. It would be necessary to replace the expression for G_t by a far more complex relation or set of relations, such as:

$$G_t = k_1 (u_{t-1})^{\rho} + k_2 (Y_{t-1} - Y_{t-2}) .$$

Even a relatively simple consideration such as this policy makes a mathematical treatment essentially unmanageable. The mathematical difficulties will depend upon the exponent ρ and furthermore an extra relation or set of relations will be needed to update the values of u_t, the level of unemployment at time t. However, the simulation of the time series generated by the new system is no more difficult than the simulation of the time series generated by the first system. In the first instance, the set of equations; the closed form expressions expressing Y_t, C_t and I_t as functions of time and the parameters; and the time series produced by a simulation can all be regarded as solutions to the problem of predicting future states of the system and the effect of changes in parameters. However these three different solution concepts are of varying usefulness to the individual who wants to answer specific questions. The first is of use only to the person of sufficient genius that he can immediately deduce the implications contained in a consistent and complete set of relations without having to "solve" them further. When an explicit expression exists for the variables of interest it usually will be a more useful solution than the time series because it is more compact and may be mathematically manipulated directly to answer questions concerning the sensitivity of parameters. For example if our solution shows that:

$$Y_t = A_1(\alpha + \sqrt{\alpha^2 - 4\beta})^t + A_2(\alpha - \sqrt{\alpha^2 - 4\beta})^t$$

we can immediately examine the derivatives $\partial Y_t/\partial\alpha$ and $\partial Y_t/\partial\beta$ to estimate the behavior of Y_t as α and β vary. The same sensitivity search could be made by simulating with the original equation system and feeding in a series of different values for α and β, then comparing the resultant time series.

In the modified set of equations it is quite likely that an explicit compact specification of the relevant variables in terms of time and the parameters does not exist. In this case the time series obtained from simulation may be our only source of insight into the dynamics of the model.

In general the sensitivity testing of large dynamic systems is difficult and expensive. It is still probably more of a scientific art than a formalized procedure. The mere fact that the techniques of simulation permit the easy generation of time series does not mean that easy exhaustive procedures exist for the exploration of the sensitivity of complex models. The combinatorics of different values for different parameters mitigates against this. The success of the search by simulation rests heavily upon the substantive knowledge of the problem and a sufficient understanding to be able to interpret the findings. One way in which this understanding is obtained is by constructing simpler models of the model or its parts where the simpler versions are tractable to mathematical, verbal or diagrammatic analysis.

The second type of mathematical model familiar to economists is one which involves simultaneous relationships, the simplest example being linear supply and demand conditions. Suppose we have supply and demand given by $s_t = a + bp_t$ and $q_t = c + dp_t$. If we assume that there is a market mechanism which brings about a price which equates supply and demand then we can find the p_t^* and the $s_t^* = q_t^*$ by solving the three equations.

A simulation, as we have already noted performs its steps in tracing out a time path by operating on all relations sequentially. Hence, in general it does not cope with simulataneous relations. However it is both easy and sometimes necessary to include simultaneous relations in a simulation to solve that part by analytic or iterative procedures and then proceed sequentially through the rest of the program.

We now discuss why it is sometimes necessary to include simultaneous relations in a simulation. It has been argued that in economic life simultaneity is not really needed because by the use of a fine enough grid on the increment of time being used the error introduced by assuming the existence of a lag will be so small that it can be safely ignored. Thus if the $\Delta t = t - (t-1)$ is small enough

$$\begin{cases} s_t = a + bp_t \\ q_t = c + dp_t \end{cases} \quad \text{or} \quad \begin{cases} s_t = a + bp_{t-1} \\ q_t = c + dp_{t-1} \end{cases}$$

will give approximately the same results. This is certainly not mathematically sound in any very general way, but for many situations of economic interest may be sufficiently true. It is almost always bound to be false however, if the time increment is large. The natural time units for socioeconomic simulations of any size are the week, month, quarter or year. In general it is very difficult to obtain large masses of data on a weekly or monthly basis and furthermore even though computers are fast there is around a 50 to 1 running time difference between a weekly or annual time basis. If, owing to considerations of the availability and quality of data, simulation time or output processing, a quarterly or yearly time scale is used then the errors introduced by replacing an equation such as $C_t = \alpha Y_t$ by $C_t = \alpha Y_{t-1}$ are large. In order to avoid such errors and still use a large time scale it is necessary to include simultaneous relations and solve them.

1.1.1. Behavior Equations or Explicit Maximization

An important distinction between different models of behavior is whether they are presented as the result of some more basic optimizing process explicitly or merely as behavior which may or may not have implicitly an explanation based on optimization. For example we may describe the behavior of a consumer by first stating that we have (say by psychometric experimentation) determined his preference system which is given by $U = - a(M-x)^2 + y$ where x and y are quantities of two goods; then observing that if the exchange rate between the two is such that $y = px$, he acts to:

$$\underset{x}{\text{Maximize}} \quad U = - a(M-x)^2 + px .$$

We might also have a behavior equation which states that x is chosen to satisfy:

$$x = M - \frac{p}{2a} \quad .$$

We could say that the latter equation implicitly represents the maximization process of the consumer, described by the former equation. The former implies the latter, but not vice-versa.

Generally a simulation will contain behavior relations and not explicit optimization processes. It is a matter of economic, political or other observations, theories and measurements to decide if the behavior equations are implied by specific optimization processes.

The policy-maker or experimenter interested in control and some criterion of optimality may easily be interested in using a simulation to explore the implications of various policies. If he is interested in optimization he can explore for an optimal policy implicitly by varying parameters or changing the structure of the policy equations, then observing the different time series generated and comparing their optimality according to some objective function or criterion outside of the simulation. For example, in the equations in 1.1. government policy was given by:

$$G_t = k \quad .$$

We could try various values of k and obtain time series for them. Our objective function might be a function of the size of the fluctuations in the business cycle.

It would then be possible to judge the worth of the different k, or of alternative policies calling for more complex government behavior.

It is, of course, possible to include an explicit maximization process in part of the program of a simulation, however the methods of simulation are used in general where the explicit maximization problem cannot necessarily be set up, and even if it can be, then it is beyond mathematical analysis. An important example of the relationship between programming and algorithm methods and simulation is provided by the various large process models of the operation of oil-refineries. The objective function is usually more or less well-defined and the processes can be described with a high degree of accuracy. Unless there are highly complex nonlinear interrelations the whole process can be described in terms of a large linear or quadratic program which may be solved by the appropriate algorithms to directly yield an optimum policy. Beyond some level of complexity however, a simulation of the process may be as much as can be done.

In one sense the exclusion of the objective function as an explicit part of a simulation is desirable. It means in particular that there is a more or less natural separation of policy considerations and the problem of describing the socioeconomic environment. This can be useful in focussing on the problems of communication between policy oriented individuals and behavioral scientists.

1.1.2. Simulation, Gaming and Game Theory

This section is included merely to clear up some misconceptions which have arisen concerning the use of three different approaches to social, political and economic problems.

We are using the word simulation here to refer to the computer "operation of a (socio-politico-economic) model. The rules of operation are specified by the behavioral and accounting equations. Uncertainty in measurement or other forms of uncertainty may be accounted for by the presence of random variables in the model. However, once the model has been formalized the insights and errors of its builder will be reflected without question by the simulation. The word simulation has been used in a somewhat different meaning [1] to refer to situations in which people play the roles of decision-makers in more or less formally simulated environments. This type of investigation is also referred to as gaming. The key difference between what we wish to refer to as simulation and as gaming is that the former involves the manipulation of a mathematical model with no direct participation of individuals; whereas the latter always involves the active participation of people as an integral part of the exercise. Simulation and gaming may easily be used as complementary investigations. For example in situations where alternative policies are being explored one can use a simulation of the variety described here to serve as the basis for a gaming exercise where instead of having explicit formulae for say, government policy and foreign trade policy two sets of players make their decisions each period, based upon the information supplied by the other parts of the model.

Gaming, especially with players simulating the roles of decision-makers has been used for teaching, training, operational and experimental purposes. Our prime concern here is with simulation, hence we terminate further discussion of gaming at this time [2].

Game theory is a third topic often confused with simulation and gaming. It is often relevant to gaming exercises or to the construction of simulations, but is by no means the same as either. The theory of games is basically a mathematical theory devoted to the study of problems of strategy, conflict and cooperation. It is implicitly relevant to the type of simulation described here inasmuch as the social, political or economic forces may reflect situations which even explicitly cannot be described in terms of maximization, but must be thought of in terms of varying levels of conflict and cooperation[3/]. However it must be stressed that the relevance of game theory is at the conceptual level inasmuch as the conflict and cooperation in the society will have already been reflected in the behavior equations.

1.2. Econometrics, Data Processing and Simulation

Simulation and econometric or sociometric methods are complementary approaches to the study of socioeconomic systems and not substitutes. The emphasis in econometrics or sociometrics is upon measurement; while the emphasis in simulation is upon control (in a feedback process). The contrast can be seen between prediction and anticipation. A problem in econometrics and production is to estimate the number of automobiles sold next year, on the assumption that certain conditions in the environment can be taken to be given. A problem in anticipation and control is to estimate the expected worth (according to some objective function) of taking an act which might influence the number of automobiles sold. The first serves as a natural data input for the second but not vice-versa.

The problems of control, planning and forecasting are all highly related but are nevertheless different. The econometrician is less concerned with the costs and value of his investigations in terms of policy, than he is with meeting various criteria of accuracy. Furthermore qua econometrician he has no interest in considering the effects of interactions between himself and the system. If he is not in a position of control this is most reasonable, as at most his statements and actions will produce effects of a second order which can be ignored.

1.2.1. Crude Data Processing and Model Building

1.2.1.1 Costs of Information

Underlying any work in planning, control or measurement in the behavioral sciences is the need for a data base and a conceptual scheme to guide the processing of the data once they have been obtained. Large scale cross-section or time series data are often needed and unless they are more or less available from the operating information of some organization accessible to the investigators, the costs of data gathering can become prohibitive both in terms of time and money.

An important joint strategy aimed at minimizing data costs is a quid pro quo arrangement between governmental and other institutions whose prime concern is with planning, and academic institutions and others who are more concerned with the development of theory and with measurement. Joint discussion and survey of data gathering requirements can result in the improvement of operating statistics coincidental with their being available at little cost for the testing of theories, econometric and sociometric investigation and other long term activities necessary to the building up of the appropriate conditions for improvements in planning.

1.2.1.2 <u>Checking Data and Models</u>

There are a few simple benefits which can be obtained from the construction of a computer simulation which are so basic that they can be easily overlooked, and yet are sufficiently valuable that by themselves they can often justify the work entailed in the simulation. They are the construction of more or less automatic procedures for the checking for logical consistency and completeness in models and the semi-automation of dimension checking and reporting unit checking on basic relations among data.

Once a system contains more than five or six variables, if no formal conceptual scheme exists which uses all of them, it is very easy for inaccuracies and inconsistencies to creep into any data-gathering scheme. A formal conceptual model <u>guards</u> against this. Possibly one of the greatest values of Leontief's input-output work has been the providing of a model which imposed order and clarity over a diverse and poorly coordinated set of data and informal concepts concerning interrelationships in the economy.

The problem of consistency in reporting units is a minor and annoying one which involves a small amount of intellectual effort, but certainly prior to computerized data handling, a great amount of work. Large data systems are often modified by changes in indexes and reporting units such as changes from long tons to short, adjustments for stock splits, changes in the definition of a reporting period or district and so forth. Unless there are formalized procedures to check through the implications of a change in one part of a reporting system, to all the other parts, within a few years the coherence in a large scale data-gathering process may be lost and it drifts apart. A by-product of a simulation when used with a data bank and processing scheme such as national economic accounts is that it serves as an automated process for checking consistency in units.

A more important problem exists in the checking of consistency in individual equations and then in the system as a whole. Before we bother with the measurement of parameters and the estimation of equations we must ask if the form of the equation is at least consistent with the basic facts we know. One very useful technique much used in physics, but hardly used in the behavioral sciences is dimension checking. The dimensions on both sides of an equation must match. An example is given below, taken from the model for the simulation which follows in Part II. The equation for social mobility which describes transfers from the category Indian to category Mestizo is given by

$$F_{3,t-1} = \alpha_{3,9}\, N_{4,t-1} \left\{ \frac{N_{3,t-1}\, N_{4,t-1}}{\left(N_{3,t-1} + N_{4,t-1}\right)^2} \right\} \left\{ \alpha_{3,10}\, \frac{Y^2_{4,t-1}}{N_{4,t-1}} + \alpha_{3,11}\, H_{r,t-1} \right\} \frac{1}{P_{t-1}}$$

where

$F_{3,t}$ = number of Indians who transfer to Mestizo in time t

$N_{3,t}$ = Mestizo population in time t

$N_{4,t}$ = Indian population in time t

$Y^2_{4,t}$ = disposable income of Indians in time t

$H_{4,t}$ = HEW expenditures on the Indians in time t

P_t = price level in time t .

Let p stand for population and m for money, then a dimension check of this equation gives:

$$p = p\left\{\frac{pp}{p^2}\right\}\left\{\frac{mp}{p} + mp\right\}\frac{1}{m}$$

which does not check and tells us that either $\alpha_{3,11}$ should have dimension $\frac{1}{p}$, or if we wish to regard it as a pure number, then we are missing a term with dimension $\frac{1}{p}$. A little reflection will show that $H_{4,t-1}/N_{4,t-1}$ is a more reasonable functional form than the one given, inasmuch as we were considering per capita effects. Having made this modification we may now observe that $\alpha_{3,9}$ (or alternatively both $\alpha_{3,10}$ and $\alpha_{3,11}$) has to have the dimension of people rather than be a pure number.

There are also logical limitations imposed on the functional form. In this case, this equation is obviously not defined for:

$$\frac{\alpha_{3,9}}{2\,P_{t-1}}\left\{\alpha_{3,10}\frac{Y^2_{4,t-1}}{N_{4,t-1}} + \alpha_{3,11}\frac{H_{4,t-1}}{N_{4,t-1}}\right\} > 1$$

because this would imply a transfer of a size larger than the Indian population. This type of limitation is hard to observe in verbal discussion.

In the example given in 1.1 with the modified government policy we do not have a complete model. This is easy to observe because if we tried to run this as a simulation we have no way of updating u_{t-1}. It is quite easy to observe this in this simple model; however given a few more equations or a verbal discussion it would be hard to observe that there were a gap or an inconsistency in the model.

Planning, data-gathering and national accounting, measurement and theory are allied but different occupations. There is a considerable gap in communication between practitioners. Methods of checking and validating which are routine to some are more or less unknown to others. The relatively simple and stringent rules which must be applied to formalizing a model for simulation can be of direct benefit in helping to bridge the gap between quantitative and qualitative approaches and between planning and research in the behavioral sciences by providing a language more precise than ordinary language and yet far less restrictive than mathematics.

1.2.1.3 Measurement, Validity of Measurement and Costs

A natural question in an econometric or sociometric study is: "how well can we measure a variable?" As we have already noted, the natural question in the context of planning and control is: "what are the costs and values of measuring this variable; and what are the gains associated with varying degrees of accuracy?"

A simulation can be regarded as a preliminary search mechanism, which in combination with the considerations of theory helps to guide the selection of aspects of the socio-economy on which the major expenditure of time and skill in measurement should be spent. This can be done by building the first models relatively quickly and crudely using whatever statistics, special knowledge, insights and theory that are easily obtainable; with this model a sensitivity search on a few variables, parameters or functional forms deemed to be important can indicate whether their accurate measurement is critical (for example, how important is the substitution of a C.E.S. function for a Cobb-Douglas function in various growth models?).

A peculiar and somewhat inverted use of simulation which can be of considerable use both in experimentation and teaching is mentioned here only to stress the interrelation between the different arts of model construction, theory and measurement. A simulation can be used as a data generating device to test the power of econometric or other methods of observing a society. A standard criticism of simulation is that although it may be quite complicated, it is still a considerable oversimplification of human affairs. A standard excuse for the failure of political, social or economic predictions is that some exogeneous event intervened or that the data were poor and hard to clean up. The output of data from a simulation provides clean statistics with no exogenous events "spoiling them"; these statistics can serve as a laboratory for testing the powers of methods of measurement, theory and insight.

1.2.2. Types of Model

Both the uses and types of simulation may be different. In physics a simulation may be used to calculate the behavior of many particles behaving according to some random process. The overall behavior of the system is predicted by many replications of a Monte Carlo process [4/]. It is possible to devise a Monte Carlo method to evaluate complicated integrals or approximate the value of π . These uses are called simulation by those who use them, however they are far removed both in intent and technique from many of the major uses in the behavorial sciences or in planning. In psychology there is considerable work on simulation of cognitive processes [5/] and this is linked with investigations into artificial intelligence [6/]; these too are of no direct concern at this time.

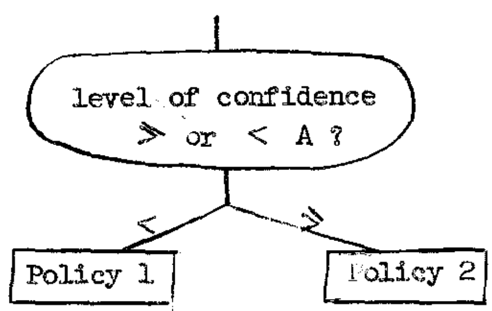
In economics there are at least four areas of application of simulation which differ both in intent and techniques. These are econometrics, planning and policy, mass data processing and economic theory. The types of consideration which distinguish the different types of simulation are:

(1) How easy is it to obtain a good representation of the system to be simulated?

(2) Can an objective function or some other clear criterion by which results can be judged, be easily or uniquely defined?

(3) How important are random elements in the simulation?

(4) Are logical switches necessary? For example if the system is to include decisions with alternatives to be selected according to criteria depending upon indices whose values may change during the simulation, they will be needed. There might, for instance be two investment policies depending upon a level of confidence:



(5) Are the number of variables and parameters so large that organization and machine problems will be critical?

(6) Is it necessary to solve simultaneous relations within the simulation?

(7) Is set searching and list processing of importance to the simulation? For example in a simulation which may contain a set of firms (say retailers) it may be desirable to locate all retailers in the Northwest who are above a certain size and who have more than k items of product 5 in stock. This involves the ability to classify special sets of firms and to search through larger sets checking for fixed or variable attributes. There is a great variation in the types of computer language and program to handle this type of activity.

(8) How important is mass data processing? There is a considerable difference between problems which involve relatively simple manipulation of large amounts of data (such as accounting schemes) and those which require the complex processing of smaller amounts.

(9) Are relatively sophisticated mathematical techniques of importance? With maximization processes, or the need to solve simultaneous systems or differential equations the mathematical computational needs are relatively high; with many business oriented models they are low.

(10) Does the model need to be flexible? Are many changes, revisions and modifications contemplated as a matter of course?. For planning, experimental and exploratory simulations, in general, the answer would be yes.

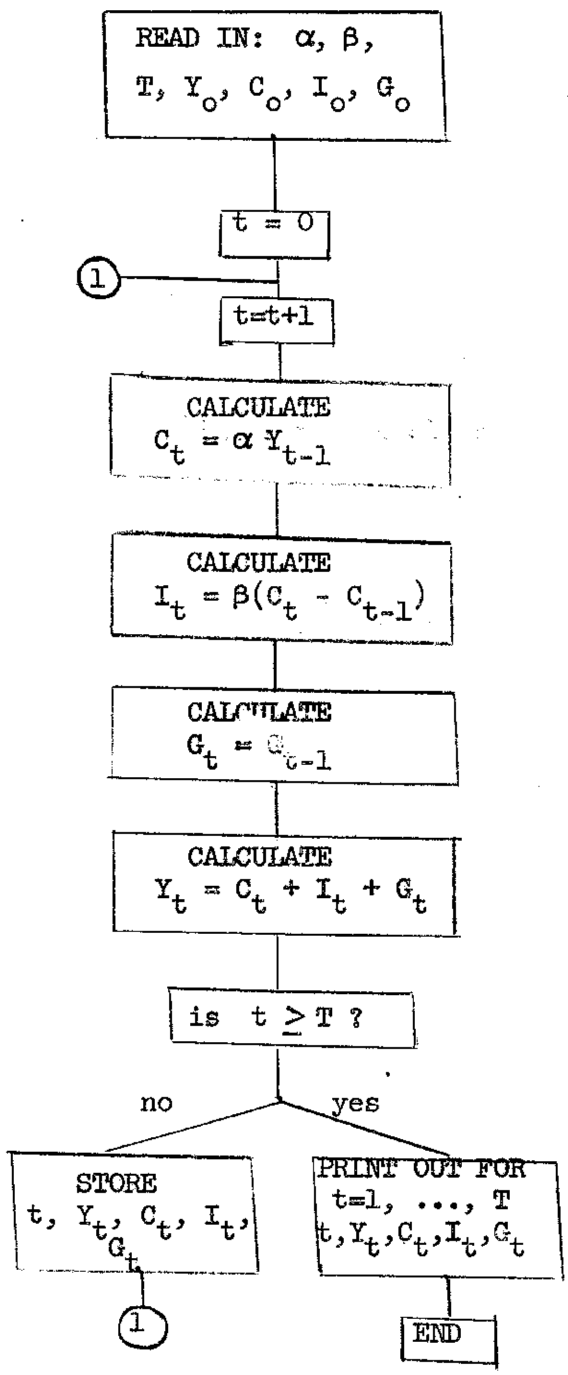
(11) Is the program intended for "on-line" use? In other words is there an immediate feedback from the user to the program calling for immediate response. Ideally in a short term planning process using a simulation, the user should be able to interrogate the computer directly. There is nevertheless a considerable difference in design between a program constructed solely for research or operations.

(12) How important is input-output and the appearance and flexibility of the format in which output appears? Do you need graphical, statistical or tabular output?

The borders between science and art, the understanding of technique and of principles have always been ill-defined. Several of the points raised above appear to be purely technical considerations. For instance, most of the problems in (12) sound as if they should be the concern of a typesetter and not a behavioral scientist. Unfortunately this is not the case. Considerations such as this are not sufficient to guarantee success but they are almost always necessary.

1.2.2.1 The Representative Unit Model

The model of Kalman Cohen of the shoe, leather and hide industries [7], the work of Forrester [8], Holland and Gillespie [9] and the Klein-Goldberger [10] and other econometric models (whose equation systems can be operated as simulations) are all examples of models using a representative behavioral unit. Whatever the level of disaggregation is in any of these models, each class is only represented by a single representative member of that class. Furthermore all of these models work on a "constant clock," in other words all actions by all actors take place at fixed increments of time; the representative consumer or investor is not late or early with his payments from period to period. A simple example of a simulation of this type can be illustrated with the difference equation model given in 1.1.



Each period the representative consumer consumes an amount proportional to previous GNP (we could introduce a random element replacing the consumption equation by $C_t = \alpha Y_{t-1} + \xi_t$) . Given the information on consumption it is now possible to calculate the aggregate investment based upon the single equation used to represent the aggregate behavior. After this government investment, then GNP can be calculated, the "clock" or the time period is then augmented by a fixed increment and the process is repeated until period T is reached, at which point the simulation stops and the time series for t, Y_t, C_t, I_t and G_t are printed out. Had C_t been related directly to I_t these two values would have had to be determined simultaneously as a block sequential or recursive model.

In resumé, we observe that this type of simulation is based on a sequential or block sequential, fixed clock model with only one member in any behavior category and with no policy or behavior alternatives, in other words, no switches.

1.2.2.2 The Cross Section Model

Examples of this type of model are provided by Orcutt 11/ and Balderston and Hoggatt 12/. In these works each behavior category is represented by a set of behavioral units or a sample. This is basically a micro-economic approach in which the behavioral conditions are based upon consideration of the individual unit rather than a representation from a statistical measure of aggregate behavior. Both of the works noted also use a fixed clock; however with many members in a behavioral category it is possible to introduce distributed time lags or a variable clock. This means that even though individuals may belong to the same behavior class, their decision times may be scattered over the time interval, a feature which may introduce considerable smoothing into otherwise highly erratic processes. The use of sets of decision-units with the same behavior lengthens simulation time considerbly and introduce difficult problems concerning sample size, as has been noted by Orcutt [13/].

1.2.2.3 Disaggregation

A well-written simulation may be regarded from one point of view as an aggregating-disaggregating device. Given current computer technology it is a relatively cheap and quick procedure to obtain outputs in more or less any combination or permutation of the original data bank used. In general the limitation will be caused by the availability of basic data. This is in general not true of noncomputerized reporting systems as the amount of labor involved in the modifying of output (even assuming one knows precisely what modifications are required) is great.

The division of labor of the behavioral sciences into sociology, anthropology, political science, economics and so forth has meant that the models constructed by members of different disciplines have either ignored the possible interactions between society and the economy or the economy and political life; or have assumed that *ceteris paribus* they can claim that these effects are implicitly accounted for. The flexibility of simulation as a modeling device makes it considerably easier to consider models which cross disciplinary boundaries. The study of development conditions appears to need an approach based upon the simulataneous consideration of at least sociological anthropological economic and political factors together.

Most economists have tended to construct disaggregated models based upon finer and finer industrial and technical breakdowns. Thus a 30 x 30 sector input-output model may be replaced by a 200 x 200 sector model; or a two-digit industry classification by a three-digit classification. It is also possible to disaggregate according to sociological or political features. Disaggregation implies a proliferation of categories, and invariably it is desirable to control this proliferation. It is feasible to limit the level of disaggregation and for the purposes of the construction of a development model select a mixture of economic, sociological and other factors to disaggregate. For example a forty economic sector model of a developing economy may not be as valuable for understanding development as a ten economic and four social sector model.

1.2.2.4 A Comparison of Models

We have already noted the difference between simulation models and ones involving explicit maximization. A few more aspects of comparison are given.

Simulation	Input-Output and Fixed Coefficient Growth Models 14/	Activity Analysis and Programming
No explicit maximizing	No explicit maximizing	Explicit maximization
Alternative processes of complexity; reached by "choice switches" in the program.	No alternative processes	Alternative processes, usually limited in complexity; reached by the maximization process.
Structural complexity and flexibility obtained at the cost of ability to analyze mathematically.	Mathematical manipulation obtained at cost of strict limits on the model.	In general mathematically more difficult and structurally more complex than I-O .

1.3. Institutional Studies, Qualitative Methods and Simulation

1.3.1. Institutional Studies and Models as an Organizing Device

The joining together of results and insights obtained from qualitative studies and quantitative studies would be an important step in recreating a political-economy. In most instances massive empirical work, observation, discussion and writing are needed before concepts can be clarified, variables separated and formal model building, quantification and measurement can take place.

Good comprehensive descriptive studies are always valuable; but by themselves are often not sufficient for a balanced understanding of a socio-economy. Ordinary language provides for much flexibility and a desirable lack of precision which prevents concepts from being overly formalized at too early a stage; however the price paid for this lack of precision can be high. In particular (as has already been noted with reference to large data gathering systems) lengthy verbal descriptions of complex multivariate systems may easily be inconsistent or incomplete. The tendency for the argument to loose coherence among different chapters of a book is large. If this book is presented as an overview of the functioning of a country, or if it presents a thesis or a theory concerning development, it should be possible to develop a simulation which at least reflects the major ideas and serves at least as an organizing and consistency checking device for the study. It must be stressed that this includes the possibility of representing many qualitative as well as quantitative effects in the simulation.

The discipline of constructing a simulation to reflect the major aspects of the observations and the theory behind an institutional study can serve a far more general and important role than providing organization and a consistency check. If there is an interest in comparative study, by designing a common output such as a national income accounting scheme it becomes possible to compare not only the books, but the more formal models and their implications to the explanation of the same set of measures.

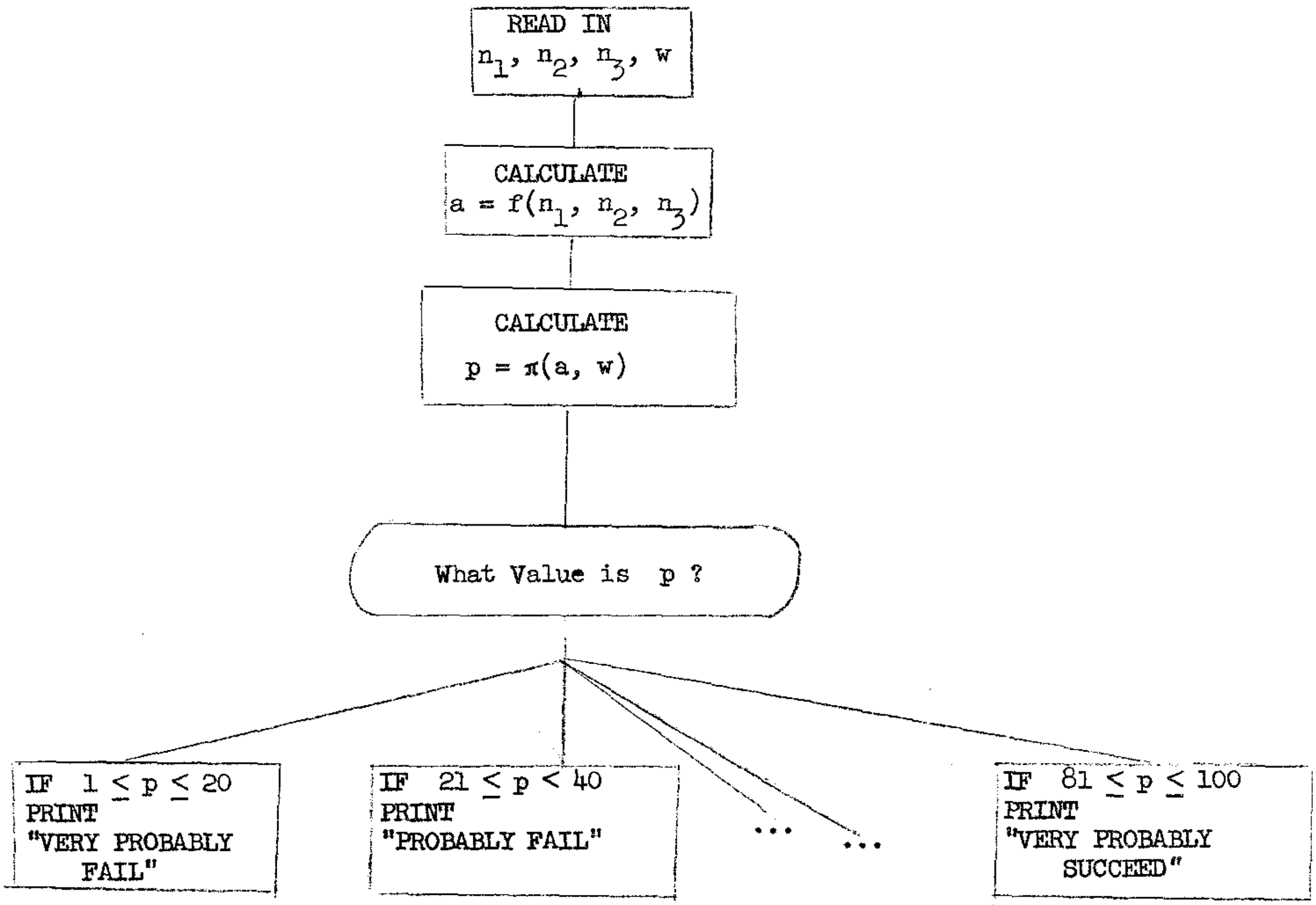
1.3.2. Qualitative Methods and Measurement

Even now there still probably exists a misunderstanding between mathematically oriented and nonmathematically oriented economists or other behavioral scientists. It is often assumed that there is a dichotomy that separates those interested in mathematical methods into a category that is willing to ignore any phenomenon to which a number cannot be attached. For many years however (at least since Edgeworth and Pareto) there has been a considerable development in methods for the study of phenomena which cannot be measured by a single number, or which need not be quantified at all. Indifference curve analysis provides one such example. More advanced methods have been concerned with the relative merits and uses of linear measures, complete orderings, partial orderings, multi-dimensional measures, lexicographic orderings [15/] and so forth. There are many phenomena involving political, social or economic choice for which the only information needed is that one is, larger, preferred to, less dangerous or more sensible than the other. By using the logical capabilities of a computer it is possible to base decisions on this type of consideration.

Work in subjective probability [16/], discernment and perception [17/] and in artificial intelligence [18/] has raised many interesting questions concerning the relation between qualitative measures, quantification of judgments and probability. By the appropriate coding, a digital computer can be used as a device to give qualitative answers. For example we consider a very simple problem. Suppose that we know that the success of Agrarian reform in a district depends upon two factors, the actual immediate wages of the farm laborers and their attitude towards the administrators sent to the district to supervise the program. It may be relatively easy to obtain a three point scale of "good," "indifferent" or "bad" for the attitude of the farm laborers from any expert; either by using a set of experts or more questions, or both we might be able to change this into say a ten point scale. From other investigations we could approximately represent the chances of success as a probability function in two variables $\pi(a, w)$ where a is the attitude measure and w the level of wages. The operations research for and the simulation of this problem would then be:

Gather a sample of expert opinion based on "good," "indifferent" or "bad." Say from n experts we have n_1 n_2 and n_3 in each category.

Select a method which gives us $a = f(n_1, n_2, n_3)$. From other research concerning the effect of wages <u>ceteris paribus</u> we have the probability of success $p(a^*) = \pi(a^*, w)$. The program is then as follows:



In this very simplified example the inputs were partially based on qualitative evaluations and the output was a qualitative interpretation of a subjective probability. This type of output may be more preferred by many than the direct statement of the probability. There are many situations where it may not be possible to use this type of quantification, however there appear to be many of interest to which this can be applied. This is especially true where the costs or possibility of the use of other methods than the interviewing of expert opinion, are prohibitive. [19]

1.3.3. The Updated Book and Comparative Library

The communication revolution brought about by the high speed digital computer is certainly at least of the same order of magnitude as the invention of the printing press. Both the concepts of book and library will have to be and are being revised.

Books on applied social and economic topics such as country studies and statistical studies are, for the most part, at this time a relatively hard to manipulate and hard to update data bank. When new information becomes available it may be highly desirable to examine it in the context of a country study and possibly to revise a book. The costs in both time and money make book revisions inordinately expensive. If, however a book has with it an associated simulation which in turn is connected with a data processing scheme and an ongoing organization in the process of creating a data bank, the possibilities for revision are considerably improved.

Specifically we consider three types of situation. The first is the standard type of updating and experimentation that both the reader and writer would like to do if the time costs were relatively low. They are, having certain statistical information updated and being able to ask a number of "what if" questions based on modifications to the hypotheses presented or on additional hypotheses. A simulator written in relation to a book is an organized precis. Clearly it will not contain the richness of description and many of the finer shades of meaning that are contained in the book; however if, for instance there is a description of investment behavior, the simulation would reflect the main features of this behavior. This being so, alternative hypotheses in the context of the overall work can be examined by changing the instructions representing investment. In many works, painstaking original data gathering is necessary and often the statistical tables presented come from nonroutine processing. Once they have been gathered, if they are important it may be possible to routinize their updating, if it is not, then obviously no simulation or formalized data handling system can help. However there is invariably a large amount of relevant data for which routinized processes can or do exist. These serve as an updating input for the simulation which in turn means that many aspects of a book remain current enlarging the basis for further work.

The second situation concerns joint work between academic research and planning groups and quantitatively and qualitatively oriented individuals. In 1.3.2. the relation between qualitative and quantitative methods has been discussed. There is little question that a simulation is no substitute for substantive knowledge and insight; it can however be an organizing device for the substantive knowledge and insights of many independent people. An interchange is needed between governmental planning groups; financial agencies and researchers in academic institutions. It is needed because the two former are in possession and control of much which the latter need; while the latter are able to take a longer term and more dispassionate view of processes and to provide insights, and a theoretical framework which helps to blend the long term with the short term operational view.

Once a model has been constructed it serves a focal point for interchange. The overall operation of it with guesses and approximations can be of operational use, while the careful measurement of technological and behavioral relations, redesign of subroutines will for the most part be of direct academic interest. Furthermore both academics and others need to establish methods or even an institution for the more effective interchange and use of information. There are large increasing returns to scale in information handling and in joint model building which have not yet been realized.

The third situation concerns the use of simulation and formal models for comparative purposes and "library cross-reference" for institutions or agencies engaged in comparative studies. If an institution such as the International Monetary Fund or a university center devoted to growth and development economics has sponsored comparative studies, simulations related to each study may be used to output in the same format for purposes of comparison. In the model presented in Part II it is for this reason that the output utilizes heavily, the National Income Accounts scheme [20/] and part of the International Monetary Fund statistics.

The area study books may be highly different both in style and content. The differences in content will be reflected in the different structures and relations of the main parts of the models. But even at this level a great potential for easy comparison of fundamental features in the socioeconomy appear. As has been noted above each model is in some sense a formal précis and hence comparison between models is easier than between books. Going back to a previous example, each model will probably have a subroutine describing investment; they may well all be different; and these differences will be easy to compare and experiment with.

Comparison at the output level can be facilitated in two ways. They are by having at least one set of output common to each model, and further by letting this output be highly related to some heavily used and agreed upon national or international set of standards. This point will be further developed in a subsequent related paper.

In resumé, any large operation to carry out comparative studies needs a basis to provide the interlinkage between field work, theory, structural studies and data gathering. The discipline and methodology of simulation may be of considerable aid in providing such a basis.

FOOTNOTES

1. See Guetzkow, H. (ed.) Simulation in Social Science. Readings. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1962.

2. For Further reference on gaming see: Proceedings of the National Symposium on Management Games, Lawrence, The University of Kansas, May, 1959.

3. For further reference on game theory see: Luce, R. D. and H. Raiffa, Games and Decisions, New York: John Wiley and Sons, Inc., 1957.

4. McCracken, D. D., "The Monte Carlo Method," Scientific American, 192.5, May, 1958.

5. Feigenbaum, E. A. and J. Feldman, Computers and Thought, pages 269-278, McGraw-Hill Book Company, 1963.

6. Minsky, M. L., "Steps Toward Artificial Intelligence," pages 8-30 in Proceedings of the IRE, 1961.

7. Cohen, K. J., Computer Models of the Shoe, Leather, Hide Sequence, Englewood Cliffs, Prentice-Hall, Inc., 1960.

8. Forrester, J. W., Industrial Dynamics, M.I.T. Press and John Wiley and Sons, 1961.

9. Holland, Edward P., with Robert W. Gillespie, Experiments on a Simulated Underdeveloped Economy: Development Plans and Balance-of-Payments Policies, M.I.T. Press, 1963.

10. Klein, L. R. and A. S. Goldberger, An Econometric Model of the United States, 1929-1952, Amsterdam: North Holland Publishing Company, 1955.

11. Orcutt, G. H., M. Greenberger, J. Korbel, and A. H. Rivlin, Microanalysis of Socioeconomic Systems - A Simulation Study, New York: Harpers, 1961.

12. Balderston, F. E. and A. C. Hoggatt, "The Simulation of Market Processes," Working Paper No. 22, Management Science Research Group, University of California, Berkeley, October, 1960.

13. Orcutt, G. H., M. Greenberger, J. Korbel, and A. H. Rivlin, op. cit.

14. See for example: von Neumann, John, "A Model of General Economic Equilibrium," in Review of Economic Studies, Vol. 13, 1945-1946.

15. Thrall, R. M., C. H. Coombs, and F. L. Davis, Decision Processes, New York: John Wiley and Sons, 1954.

16. Pratt, John W., Howard Raiffa, and Robert Schlaifer, "The Foundation of Decision Under Uncertainty: An Elementary Exposition," pages 353-375 in Journal of the American Statistical Association, Vol. 59, No. 306, June, 1964.

17. Miller, George A., "The Magical Number Seven, Plus or Minus Two -- Some Limits on our Capacity for Processing Information," Journal Psychological Review, Vol. 63, No. 2, 1956.

18. See Feigenbaum, E. A. and J. Feldman, op. cit.

19. See an interesting study: Gordon, T. J., and Olaf Helmer, Report on a Long-Range Forecasting Study, The RAND Corporation, September, 1964.

20. Abraham, William, Nancy D. Ruggles, and Richard Ruggles, An Economic Data Reporting System for the Agency for International Development, New Haven, Connecticut: Economic Growth Center at Yale University, 1965.